

Datenformat zum Import von CSV-Dateien (Eingabe für das BJ 2019, Stand September 2019)

Allgemeines

Zur Vereinfachung der Dateneingabe für die Deutsche Bibliotheksstatistik (DBS) haben die **Fachstellen** die Möglichkeit, die Daten der von ihnen betreuten Bibliotheken direkt zu importieren. Die Dateien im **CSV-Format** (*Comma Separated Values*)¹ werden dabei über ein Web-Formular hochgeladen, geprüft und nach erfolgreicher Kontrolle in die Datenbank der DBS eingespielt. Dafür muss für jede Fachstelle (genauer: jede Fachstellen-ID) jeweils genau eine Import-Datei geliefert werden. Unmittelbar danach kann auf die importierten Daten zugegriffen werden.

Dieser Weg zum Import der DBS-Daten richtet sich ausschließlich an Fachstellen!!!

Format

Grundsätzlich erzeugt **Microsoft Excel** Export-Dateien in einem für den DBS-Import gültigen Format². Es sind dabei folgende Hinweise zu beachten:

1. Erwartet werden Dateien im **ASCII-Format**, die Zeichensatz-Kodierung sollte „Western ISO-8859-1“ entsprechen (also kein Unicode / UTF8 o.ä.).
2. Jede Zeile enthält mehrere Spalten, die durch **Semikolon** (»;«) voneinander getrennt sind. Die Werte der einzelnen Spalten können durch doppelte oder einfache Anführungszeichen (»"«, »'«) eingeschlossen sein, notwendig ist dies jedoch nicht.
3. Die erste Zeile kann **Überschriften** enthalten, dies ist aber nicht erforderlich. Leerzeilen werden beim Import übersprungen.
4. Für den Import der Daten sind in jeder Zeile (mindestens) 96 **Spalten** notwendig, das heißt eine Spalte für die DBS-ID und 95 Spalten für die Daten des Grundfragebogens („Fragebogen für Öffentliche Bibliotheken“). Zum Format der Spalten:
 - a. In der ersten Spalte steht die **DBS-Identifikationsnummer** (DBS-ID), die jeweils aus genau fünf Zeichen besteht, zwei Buchstaben gefolgt von drei Ziffern (z.B. „AB123“).
 - b. Die übrigen Spalten enthalten die **Antworten** des Fragebogens. Dabei gibt es aktuell drei verschiedene Datentypen:
 - i. **Numerisch, ohne Nachkommastelle**: Hier wird jeweils eine (maximal 11-stellige) ganze Zahl erwartet, Nachkommastellen werden nicht berücksichtigt, Tausenderpunkte sollten nicht enthalten sein.

¹ vgl. Beispiel auf Seite 3

² unter Office 2000: Menüpunkt „Datei“: „Speichern unter...“, dann im Dialog die Auswahlliste „Dateityp“: „CSV (Trennzeichen getrennt) (*.csv)“ wählen

- ii. **Numerisch, zwei Nachkommastellen:** Wie unter i., jedoch können bis zu zwei Nachkommastellen mit Komma getrennt angegeben werden (weitere Stellen werden ignoriert).
 - iii. **Auswahlliste:** Dabei kann im Fragebogen per „Drop-Down-Box“ jeweils eine von mehreren Alternativen ausgewählt werden (meist „ja/nein“-Fragen). **Hier darf die Datei nur noch den Klartext** (exakt wie im Fragebogen, keine Abkürzung) **enthalten**.³
- c. Wichtig ist der Unterschied zwischen den Einträgen „**N**“ und „**0**“ (**Null**): Ein „N“ besagt „vorhanden, aber nicht erfasst“ (keine Zählung möglich, Wert unbekannt), eine Null hingegen bedeutet „der Wert beträgt Null“. Bei Auswahllisten kann das Feld auch leer bleiben.
- d. Die **Zahl der Spalten** ist variabel, je nachdem, welche Teile des Fragenkanons importiert werden sollen. Es gibt genau vier Möglichkeiten:
- i. **96 Spalten:** Damit wird ausschließlich der Hauptteil des Fragebogens importiert (Fragen 1-104 und Frage 199).
 - ii. **107 Spalten:** Dann werden der Hauptteil und der Abschnitt für Patientenbibliotheken (Fragengruppe 200, Fragen 200-210) importiert.
 - iii. **114 Spalten:** Importiert wird wie unter (ii.). Hinzu kommt der Abschnitt für Fahrbibliotheken (Fragengruppe 300, Fragen 300-307). **Die Fragen der zusätzlichen Fahrbibliothekenumfrage werden nicht importiert.**
 - iv. **136 Spalten:** Hier wird wie unter (iii.) importiert. Hinzu kommt der Abschnitt für Musikbibliotheken (Fragengruppe 400, Fragen 400 – 421)

Wichtig ist, dass jede Zeile der Import-Datei jeweils genau die oben angegebene Zahl von Spalten enthält.

5. **Bereits existierende Daten**, die von Bibliotheken evtl. schon von Hand eingegeben wurden, werden beim Import auf jeden Fall überschrieben. Dies gilt jedoch nur für die Bereiche, die im Import enthalten sind. Bei importierten 96 Spalten z.B. bleiben die Fragen 200-210 etc. unberührt.

Tabelle 2 zeigt, welche csv-Spalte der Frage des DBS-Fragebogens entspricht.

Prüfungen

Beim Import werden verschiedene Prüfungen durchgeführt, die das Format und teilweise auch den Inhalt der Daten betreffen.

1. Es wird geprüft, ob die **DBS-ID** in der ersten Spalte das korrekte Format hat („AB123“), ob sie in der DBS existiert und ob die beim Import eingestellte Fachstelle überhaupt für diese Bibliothek zuständig ist (dies ist den DBS-Adressdaten zu entnehmen).

³ Beispiel: Bei einem Auswahl-Feld kann diese Spalte beim Import also entweder leer bleiben (= „nicht erfasst“) oder die Werte „ja“ oder „nein“ oder „teilweise“ (nur bei Frage 38.2) enthalten.

2. **Numerische Felder** werden auf das korrekte Format hin untersucht (Zahlen mit/ohne Nachkommastellen). Bei Feldern mit **Auswahllisten** wird geprüft, ob der Inhalt tatsächlich einer möglichen Option entspricht.

Beispiel

Hier nun ein Beispiel für den Aufbau einer Import-Datei (in tabellarischer Form etwa wie in Excel):

Tabelle 1:

Spalte1 DBS-ID	2 primäre Nutzer	3 Zahl Bibl.	4 Zweig- stel- len	...	59 Ei- gene Ein- nah- men	60 Ben.- geb.	...	95 Mobile Endge- räte.	96 Anmer- kungen
Frage- Nr. (*)	1	2	3	...	63	65	...	104	199
AB123	13538	1	0	...	12900	nein	...	ja	
CD456	9877	1	0	...	5890		...	nein	1 Monat ge- schlos- sen
EF789	25987	2	1	...	23980	ja	...		

(*) Die Frage-Nr. dient hier nur als Orientierungshilfe und ist nicht Bestandteil einer Import-Datei

In der Import-Datei im **CSV-Format** würde dies dann z.B. so aussehen:

```
AB123;13538;1;0;...;12900;nein;...;2;;
CD456;9877;1;0;...;5890;;...;3;1 Monat geschlossen;
EF789;25987;2;1;...;23980;ja;...;1;;
```

Hinweis für den Export aus Microsoft Excel

Beim Export aus Excel heraus ist zu beachten, dass **ausgeblendete Spalten** (oder Zeilen) immer komplett mit exportiert werden. Hier ist die folgende Technik zu empfehlen: Markieren Sie in der Excel-Tabelle alle Zellen, die zum Import gehören sollen, kopieren Sie diese Zellen und fügen Sie diese auf einem neuen Blatt über „Inhalte einfügen...“ (im Menü „Bearbeiten“) ein, wobei Sie im angezeigten Dialog „Einfügen: Werte“ auswählen sollten. Auf diesem Weg können Sie in der ursprünglichen Tabelle auch berechnete Zellen (Formeln) beibehalten.

Tabelle 2: Konkordanz csv-Spalte – DBS-Frage

<i>csv-Spalte</i>		
1	DBS-ID	
	DBS-Frage	Gruppe
2	1	1
3	2	1
4	3	1
5	4	1
6	5	1
7	6	1
8	6.1	1
9	7	1
10	7.1	1
11	8	1
12	8.1	1
13	9	2
14	10.1	2
15	10.2	2
16	11	2
17	12	2
18	12.1	2
19	13	3
20	14	3
21	14.1	3
22	15	3
23	16	3
24	17	3
25	18	3
26	19	3
27	24	3
28	25	3
29	28	3
30	29	3
31	34	3
32	34.1	3
33	34.2	3
34	35	3
35	36	3
36	38	3
37	38.1	3
38	38.2	3
39	39	3
40	40	3
41	41.1	3
42	43	3
43	44	3
44	49	4

DBS - Datenformat zum Import von CSV-Dateien

<i>csv-Spalte</i>		
45	50	4
46	50.1	4
47	51	4
48	52	4
49	53	4
50	54	4
51	55	4
52	56	4
53	57	4
54	58	4
55	59	4
56	60	4
57	61	4
58	62	4
59	63	4
60	65	4
61	66	5
62	67	5
63	68	5
64	69	5
65	70	5
66	72	5
67	74	5
68	75	5
69	76	5
70	77	5
71	78	5
72	80	6
73	81	6
74	82	6
75	83	6
76	85	6
77	86	6
78	87	6
79	88	6
80	89	6
81	90	6
82	91	6
83	92	6
84	94	6
85	95	6
86	96	6
87	97	6
88	98	6
89	99	6
90	99.1	6

DBS - Datenformat zum Import von CSV-Dateien

<i>csv-Spalte</i>		
91	100	6
92	101	6
93	102	6
94	103	6
95	104	6
96	199	6
97	200	7
98	201	7
99	202	7
100	203	7
101	204	7
102	205	7
103	206	7
104	207	7
105	208	7
106	209	7
107	210	7
108	300	8
109	301	8
110	302	8
111	303	8
112	304	8
113	305	8
114	307	8
115	400	9
116	401	9
117	402	9
118	403	9
119	404	9
120	405	9
121	406	9
122	407	9
123	408	9
124	409	9
125	410	9
126	411	9
127	412	9
128	413	9
129	414	9
130	415	9
131	416	9
132	417	9
133	418	9
134	419	9
135	420	9
136	421	9

DBS - Datenformat zum Import von CSV-Dateien

Kontakt per E-Mail: dbs@hbz-nrw.de

Homepage: <https://www.hbz-nrw.de/produkte/bibliotheksstatistik>

Stand: 04.11.2019